UNIVERSITY OF MINNESOTA

This is to certify that I have examined this copy of a Plan B Project by

Melinda Marie Dertinger

and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the final examining committee have been made.

Aydin Durgunoglu, PhD

Name of Faculty Advisor

ADurgunoglu

Signature of Faculty Advisor

5/30/19

Date

Using Counter-Stereotypical Exemplars as an Intervention to Reduce Implicit Gender Stereotypes: An Application of Gender Schema Theory

A THESIS SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

Melinda Marie Dertinger

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

Aydin Durgunoglu

May 2019

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor, Dr. Aydin Durgunoglu, first for seeing my potential as an undergraduate student and encouraging me to apply to graduate school, and second for continuing your support during my graduate journey. Thank you for all you have done to help me reach my goals. I would not have been able to complete this thesis without your continued guidance, support, patience, and wisdom. I am honored to have been guided by such a wise and compassionate person.

I would also like to thank my family for all of their support and encouragement during the entirety of my graduate education. Thank you for always staying by my side and providing kind and motivating words during my many late nights of studying, moments of stress and exhaustion, and times of need. To my husband, Chris, thank you for being a loving, supportive, and equal partner throughout this experience. To my son, Alec, thank you for motivating me and giving me reason to complete my degree, serve as a positive role model, and work to strive for and teach the importance of equality.

Finally, I would like to thank my committee members, Dr. Melissa Walls and Dr. Ashley Thompson, for the support and opportunities they provided me throughout my graduate experience. I am thankful for the opportunities you have given me during my time here to work with you on your research projects and learn so much from you in the process. Your outstanding work and contributions to your fields continues to inspire me to use the skills I've learned to try to advance scientific knowledge in a meaningful way.

Table of Contents

List of Tables

List of Figures

Key terms

- **Stereotype:** A widely held belief regarding a trait, role, characteristic, etc. about a group of people that shapes our judgments of individuals in those groups. Stereotypes are not always accurate or based in reality.
- **Stereotype threat:** According to Steele and Aronson (1995), stereotype threat is feeling "at-risk of conforming, as self-characteristic, a negative stereotype about one's group". In terms of gender, stereotype threat occurs when women are aware of existing stereotypes regarding their gender and are concerned with being judged based on this stereotype.
- **Implicit measures:** Assessments designed to evaluate thoughts, attitudes, and/or feelings outside of conscious awareness (e.g., reaction times, the Implicit Association Test; Lai et al., 2014).
- **Explicit measures:** Assessments designed to evaluate thoughts, attitudes, behaviors, and/or feelings in which participants are consciously aware (e.g., self-reports, questionnaires).
- **Counter-stereotypical images:** In the current study, counter-stereotypical images are pictures of men or women performing jobs/duties that are seen as contrary to roles typically performed by a particular gender (e.g., a male make-up artist or a female bricklayer).
- **Schema:** A mental structure that organizes categories and guides perception of information. Individuals interpret information in terms of cognitive schemas to

organize the world and provide a framework for understanding our surrounding and our own self-image (Bem, 1981).

- **Deese-Roediger-McDermott (DRM) Paradigm:** Named for its developers, the DRM paradigm is a cognitive task that measures false memories created by implicit associations or biases (Roediger & McDermott, 1995). This task presents participants with a list of related words (e.g., *bed*, *rest*, *awake*, *dream*) and later asks participants to recall the words presented; participants are likely to falsely recognize a non-presented related word (e.g., *sleep*), called a false alarm. The number of false alarms represents the rate of direct associations.
- **Lure:** In the DRM Paradigm task, a critical lure is a word in the recognition test that was not presented in the original word lists during the exposure phase but is semantically related to presented words. For example, if *bed*, *rest*, and *awake* were presented during the exposure phase, a different but related word, like *sleep,* presented in the recognition test would be a lure.
- **False alarm:** In the DRM Paradigm task, a false alarm is falsely recalling a new word (a word not presented in the exposure phase) as old; e.g., falsely recalling that *sleep* was presented during the exposure phase. For the current study, the number of false alarms recalled by each participant is also referred to as the rate of false memories.
- **Direct associations:** Words that are directly related without a mediating concept (e.g., *sleep, bed,* and *rest*).

- **Indirect associations:** Words that are indirectly related through a mediating concept (e.g., *nurse* and *librarian* are associated through their mediating link, *feminine roles*).

Abstract

Stereotypes influence the way we interpret our social environments and our own self-concept. Because many of the stereotypes we hold regarding gender are implicit, they are often difficult to overcome and contribute to systematic discrimination beginning in childhood. Gender schema theory provides a conceptual framework to guide our understanding of how we interpret information regarding others (the outward expression of stereotypes) and ourselves (the inward expression of stereotypes) based on gender. The current research seeks to examine the effects of an image-viewing manipulation on both the outward and inward expression of stereotypes. We investigated if viewing either stereotypical or counter-stereotypical gender role images influenced the number of false stereotypical memories recalled (Study 2) and the vulnerability to stereotype threat (Study 3). Overall, our hypotheses were not supported, demonstrating a lack of replicability of previous findings and suggesting that this image manipulation may not be effective for reducing implicit gender stereotypes.

Using Counter-Stereotypical Exemplars as an Intervention to Reduce Implicit Gender Stereotypes: An Application of Gender Schema Theory

Categorical Thinking & Gender Stereotypes

Humans have developed many beneficial cognitive processes that allow us to learn about, understand, and make predictions of our environments (Macrae & Bodenhausen, 2000). With limited cognitive resources available to us, one of the processes we employ involves the use of categorical thinking of other individuals. Perceiving individuals in terms of social categories has allowed us to simplify the person perception process by instantly gaining information about individuals based on the categories to which we assign them automatically (Macrae & Bodenhausen, 2000). Today, however, the benefits of categorical thinking are often overshadowed by the incorrect assumptions we make about categories of people. Specifically, categorical thinking can automatically influence our perception of others as we often use schemas to guide information processing and form impressions of individuals, which frequently results in stereotype-based judgments (Kawakami, Moll, Hermsen, Dovidio, & Russin, 2000; Macrae & Bodenhausen, 2000). Assigning an individual to a category enables us to make stereotypical assumptions of appropriate social roles and traits for that category and informs us of what we should expect from the individual based on those assumptions. For example, gender stereotypes imply assumptions of appropriate social roles for men and women, which is why we expect to see surgeons who are men and nurses who are women (Finnegan, Oakhill, & Garnham, 2015). The stereotypes that we hold today are pervasive because of the social and cognitive adaptations associated with them, and the potential payoffs for these cognitive shortcuts help explain why stereotypes are so difficult to overcome in our current social environment (Blair, 2002; Macrae & Bodenhausen, 2000).

Stereotypes penetrate many aspects of our daily social lives, and often create negative perceptions and expectations of others that can lead to systematic discrimination and inequality beginning at an early age. For example, the English language contains many role nouns that are not inherently gendered in definition, but have inherited gender-biased connotations, e.g. mechanic, plumber, secretary, and beautician (Finnegan, Oakhill, & Garnham, 2015). Our gender role expectations are so automatic that we may not even be aware of stereotype activation when we read role nouns, which perpetuates gender role stereotypes in the English language (Finnegan et al., 2015). This is problematic because such occupational stereotypes may influence career choice and job and activity preference as early as age 6, which may lead to inequality by limiting choices offered to each gender (Finnegan at al., 2015; Liben, Bigler, & Krogh, 2002), contributing to the systematic oppression of women.

Evidence for the real-world implications of gender stereotypes exist in statistics regarding education, income, and career placement for women and men, and indeed show a pattern of inequality beginning with young children. More males than females make up prekindergarten students, physics students, and advanced placement (AP) mathematics students (U.S. Department of Education, 2012). In the workforce, men earn more than women in every country with available data, with some countries holding a gender pay gap of over 20% (Workplace Gender Equality Agency, 2016). This income gap is even more concerning considering that women tend to earn a majority of higher education degrees in most countries (U.S. Department of Education, 2012; WGEA, 2016). In the U.S. specifically, the U.S. Department of Labor, Women's Bureau (2014) reports that women earn less than men at all levels of educational attainment, at all ages in life, and in all types of occupations (even including careers that are stereotypically seen as feminine, such as education, community and social service, and personal care occupations).

Of particular concern is women's underrepresentation in science, technology, engineering, and math (STEM) fields. According to the U.S. Department of Commerce (2011), although women hold around half of all jobs in the U.S., less than 25% of those in STEM fields are held by women, and only 14% of engineering jobs specifically are held by women. More men than women hold STEM jobs at every level of education, and women in STEM fields earn 14% less than men in the same fields – less than the overall gender wage gap, but still a significant inequality (U.S. Department of Commerce, 2011). Even among STEM fields, women are more likely than men to work in physical and life sciences than math, computers, or engineering, and men are more likely than women to hold STEM management positions (U.S. Department of Commerce, 2011).

Together, these data demonstrate that even though women earn more higher education degrees overall than men, women are disproportionately underrepresented in many areas – particularly STEM fields – and earn less than their male counterparts in the same fields. One factor that likely contributes to such disparities in education, occupation, and income is our widely held – often implicit – beliefs of attributes that are appropriate for each gender. These beliefs seem to be a result of categorical thinking and gender stereotypes, and may be explained in part by gender schema theory.

Gender Schema Theory & Stereotypes

Gender schema theory, developed by Sandra Bem in 1981, explains how individuals become gendered in society. A schema is a mental structure that organizes categories and guides perception of information (Bem, 1981). According to Bem (1981), individuals organize information in terms of cognitive schemas to organize the world and provide a framework for understanding our surroundings and our own self-image. Our society places much importance on a gender dichotomy (i.e., men and women), and children begin observing and learning from a very young age the appropriate traits and behaviors associated with each gender. These expectations are cognitively grouped into associations (schemas), which organize our understanding of what it means to be a man or a woman and masculine or feminine (Bem, 1981). Gender schemas also guides our perception and interpretation of new information, in that incoming information interacts with our existing schemas during the perception process (Bem, 1981).

These gender schemas not only provide information regarding roles, traits, and behaviors that we should expect to see in others, but they are also ingrained in our own self-concept as they provide information about which attributes should be linked to our gender and thus ourselves (Bem, 1981). For example, young girls are rarely praised for being strong, and young boys are rarely praised for being nurturing, which teaches children to internalize gender-specific attributes. Moreover, children evaluate their adequacy as an individual in terms of how they represent these gender schemas (Bem, 1981).

Because schemas guide our processing of incoming information as it relates to others and ourselves, gender schema theory provides an excellent framework for our understanding of both the outward and inward expression of gender stereotypes and guided our research questions in the current study. In fact, stereotypes may be thought of as “the functional equivalent of schemata”, as schemas are utilized to guide social perception and memory regarding groups of people and generalized beliefs concerning behaviors and attributes of these groups (Hudak, 1993). Accordingly, these schemas and stereotypic beliefs often result in biased associations among specific characteristics and certain groups of people that are not accurate for individuals (Hudak, 1993). These biased associations are so strong that they may even create false memories involving gender-stereotypic behavior and attributes.

False Memories, Implicit Associations, & the DRM Paradigm

The Deese-Roediger-McDermott (DRM) paradigm, named for its developers, is a cognitive task that measures false memories created by implicit associations or biases (Roediger & McDermott, 1995). Specifically, the DRM paradigm presents participants with a list of related words (e.g., *bed*, *rest*, *awake*, *dream*) and later asks participants to recognize which words were presented; participants are likely to falsely recognize a non-represented related word (e.g., *sleep*), called a false alarm (Roediger, McDermott, 1995). The number of false alarms represents the rate of direct associations – *bed* and *rest* are directly associated to *sleep* through cognitive associative networks because they are related directly without a mediating concept (Lenton, Blair, & Hastie, 2000). In other words, the presented words invoke a particular schema (a *sleep* schema) which results in a tendency to recall other similar (but not presented) words that are related, and at a rate that is similar to words actually presented.

In their study, Lenton, Blair, & Hastie (2000) took the DRM paradigm one step further and created additional lists to measure indirect associations as well as direct associations. The authors utilized some of the original lists created by Roediger and McDermott to measure direct associations, but they also used two additional lists to measure indirect gender stereotype associations. Indirect associations consist of concepts that are “related through a third, mediating concept” (Lenton et al., 2000), and are very common in social stereotypes. For example, *nurse* and *librarian* are indirectly associated through their mediating link, *feminine roles* (Lenton et al., 2000). The results of their study demonstrated rates of false memories of both direct associations and stereotype-consistent roles and traits (indirect associations) that were similar to rates of accurate memories (Lenton et al., 2000). Their findings also support gender schema theory, as it appears as though participants utilized a *feminine roles* schema while processing and remembering the words. Interestingly, however, most participants did not mention using any schemas during memory and retrieval and were not explicitly aware that the role terms were presented based on stereotypic schemas (Lenton et al., 2000). This suggests that schemas may operate outside of our conscious awareness, resulting in implicit stereotypes and biases. Because this adapted DRM paradigm demonstrated implicit gender associations in previous work, we chose it as our measure of implicit bias for the current study.

Turning Stereotypes Inward

As noted earlier, stereotypes and schemas not only influence our perception of and behavior towards others, but they also affect our own self-concepts and ability. Over the years, research regarding stereotype threat has shown that the stereotypes that we hold regarding our own gender or ethnicity can negatively influence our performance on a variety of tasks (Steele & Aronson, 1995; Spencer, Steele, & Quinn, 1999; McGlone, Aronson, & Kobrynowicz, 2006; Campbell & Collaer, 2009; Fogliati & Bussey, 2013; Neuburger, Ruthsatz, Jansen, & Quaiser-Pohl, 2015). For example, in their classic study, Steele and Aronson (1995) experimentally manipulated the occurrence of self-evaluative threat among African-American students by varying the diagnostic level of a difficult verbal examination. In other words, the authors told both African-American and European-American participants that the exam was either indicative or not indicative of their intellectual ability. They found that telling participants that the test was diagnostic of their intellectual ability significantly decreased performance of African-American students compared to European-American students, but, more importantly, telling participants that the test was not diagnostic of their overall ability resulted in no differences in performance between the two groups (Steele & Aronson, 1995). Steele and Aronson concluded that stating that a difficult test is indicative of intellectual ability invoked a negative stereotype among the African-American students that caused an anxiety of confirming the stereotype, which undermined performance (Steele and Aronson, 1995).

Since this classic study, stereotype threat has been applied to explain discrepancies in many stigmatized groups, including various ethnicities, ages, and genders (see Spencer, Logel, & Davies, 2016, for a review). In terms of gender specifically, Spencer, Steele, and Quinn (1999) found that telling participants that women typically underperform on a difficult math test evoked stereotype threat in women and resulted in reduced performance on the test. More recently, stereotype threat has been applied to many different areas involving gender, including women's performance on standardized tests (Good, Aronson, & Inzlicht, 2003), on visuospatial tests (Campbell & Collaer, 2009), in statistics courses (Franceschini, Galli, Chiese, & Primi, 2014), women's identity as research scientists (Smith, Brown, Thoman, & Deemer, 2015), women's well-being in the workplace (von Hippel, Sekaquaptewa, & McFarlane, 2015) and the gap in men and women's political knowledge (McGlone, Aronson, & Kobynowicz, 2006). Moreover, stereotypes begin negatively influencing women's academic performance in some tasks by fourth grade (Nueburger, Ruthsatz, Janse, & Quaiser-Pohl, 2015). The robust stereotype threat findings previously reported and the potential for stereotype threat to have detrimental outcomes for individuals in minority groups inspired us to include stereotype threat as a measure of the more inward, personal expression of implicit bias in the current study.

Together, the vast literature regarding stereotype threat research demonstrates that the stereotypes we hold regarding groups of people are able to influence performance and motivation on a variety of tasks. Furthermore, these stereotype effects can influence women's opportunities and success at many stages of life and in many important real-world situations, such as academic performance, career choice, workplace well-being, and income. And because our stereotypes are often implicit, we are frequently not even aware of our biases, which makes them difficult to measure and control. Nevertheless, much research has been conducted in an attempt to find interventions aimed at measuring and reducing these implicit stereotypes.

Measuring Stereotypes

Numerous different ways of measuring both the inward and outward expression of stereotypes have been examined over recent years, each with their own advantages and shortcomings. In terms of the expression of stereotypes toward others, while explicit measures are able to assess conscious biases and attitudes, they cannot provide information regarding implicit stereotypes that are beyond conscious control and are subject to response biases since responses are slow, intentional, and motivated (Akrami & Ekehammar, 2005). Conversely, implicit measures are able to assess fast, latent, and automatic stereotypes and are not subject to response bias (Nosek, Greenwald, & Banaji, 2005). One example of an implicit association measure is the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998). The IAT records accuracy and reaction time of responses from participants' pairings of concepts and attributes. As an example, participants could be asked to pair either the word *black* or *white* with words like *pleasant* or *unpleasant*. Faster reaction times and higher accuracy of responses represents stronger associations in memory, or an implicit bias (Greenwald et al., 1998).

Another method of measuring implicit bias is the DRM paradigm (discussed above). This paradigm is able to measure the external expression of implicit biases differently than the IAT. Specifically, it measures the rate of false memories created from cognitive associations. Lenton, Blair, and Hastie's (2000) adapted lists are able to measure implicit gender associations, represented by false stereotypical memories, which is why they were chosen for the current study. We hypothesized that our results would replicate those of Lenton, Blair, and Hastie (2000), but with a different measure if implicit bias (i.e., the DRM paradigm).

In terms of the inward expression of stereotypes, stereotype threat has been measured several ways. Many researchers examine this phenomenon by comparing performance on standardized math tests (Good et al., 2003; Finnigan & Corker, 2016), but other measures have also been used, including actual achievement in math courses (Good et al., 2008) and interest, enjoyment, and confidence in certain fields (Spencer et al., 2016). Overall, the literature shows that stereotype threat can be measured in a variety of situations and experiences but is most commonly and easily measured in the lab using scores on standardized tests, such as the ACT (American College Testing), SAT (Scholastic Assessment Test), or GRE (Graduate Record Examination) (Spencer et al., 2016). Thus, we selected standardized test questions to measure stereotype threat in our sample.

Overcoming Outward & Inward Stereotypes

Much research has also been conducted in an attempt to find interventions aimed at reducing stereotypes. Some of these interventions showing some promise in buffering the negative effects of stereotypes include imagining intergroup contact (Brambilla, Ravenna, & Hewstone, 2012), training to negate stereotypes (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000), exposure to counter-stereotypical exemplars (Lai et al., 2014; Finnegan et al., 2015), vividly imagining a counter-stereotypical scenario (Lai et al., 2014; Blair, Ma, & Lenton, 2001), and perspective taking (Todd, Bodenhausen, Richeson, & Galinsky, 2011). Some of these potentially effective interventions involve rather simple manipulations. For example, Finnegan et al. (2015) presented pictures of individuals working in either stereotypical or counter-stereotypical roles and found that participants exposed to counter-stereotypical images had higher accuracy in judging stereotype incongruent word pairs than those exposed to stereotypical images. In other words, participants who were exposed to counter-stereotypical images such as a male makeup artists or female bricklayer were better able to process counter-stereotypical words pairs like *brother-nurse*. The results of such research suggest that implicit and automatic stereotypes may actually be malleable. In fact, Blair (2002) suggests that automatic stereotypes may be far less inflexible than previously assumed, and they may be moderated by social motives, strategic efforts, focus of attention, and contextual cues. Importantly, research on stereotype intervention strategies suggests that interventions are successful across different domains, so a technique that reduces one type of bias will likely be successful for other biases (Finnegan et al., 2015).

One type of intervention that does not seem to be as successful in reducing the activation of stereotypes is stereotype suppression. In fact, stereotype suppression may generate ironic consequences; attempting to explicitly suppress stereotypes may actually lead to an increase in stereotype activation and application, which may lead to an increase in discriminatory behavior (Macrae & Bodenhausen, 2000). Furthermore, explicit control of overt behavior (i.e., application of stereotypes) requires important cognitive resources such as awareness of the stereotype and motivation to control discriminant behavior, resources that can be cognitively demanding (Finnegan et al., 2015). The ironic potential for stereotype rebound effects after suppression techniques suggests that this may not be the best method to reduce negative stereotype activation or application. Instead, it may be more ideal to reduce the initial activation of stereotypes (i.e., reduce the amount of negative implicit associations) rather than attempting to control the subsequent application of the stereotypes (Finnegan et al., 2015).

When it comes to stereotype threat, informing women about the phenomenon is simply not enough to buffer the negative effects in performance (Tomasetto & Appoloni, 2013). This suggests that an individual-based approach may not be effective for countering the effects of stereotype threat. In their study, Shaffer, Marx, and Prislin (2013) found that, when presenting women with information that positively highlighted the success and number of women in STEM fields, women performed as well as men on a math test, suggesting the importance of highlighting women STEM role models. Furthermore, when in conditions under threat (such as telling female participants they will be completing a diagnostic math test soon), women may use media to help counter the effects of the stereotype threat on performance (Luong & Knobloch-Westerwick, 2017), which also implies that providing positive female role models in the media may improve women's academic performance. The results of studies like these provide evidence for the importance of highlighting female STEM role models at a community or societal scope, as compared to individual-based interventions. Thus, the current study sought to investigate the effectiveness of highlighting counter-stereotypical exemplars on reducing implicit bias.

Current Research

Although many studies have explored stereotypes and stereotype threat using different measures and manipulations, to our knowledge, no research has yet to combine implicit measures of stereotypes with explicit measures of stereotype threat. Moreover, we are not aware of any research aimed at using the presentation of counter-stereotypical exemplar images as an intervention for reducing stereotype threat – much of the existing research entails informing participants that a particular test will be diagnostic of their ability or the gender differences typically seen in a test.

Additionally, as far as we know, an image intervention has not yet been tested with the implicit DRM paradigm measure. Finnegan, Oakhill, and Garnham (2015) used such an intervention by presenting either stereotypical or counter-stereotypical gender role images and then measuring accuracy and reaction time of responses when asked if a role term (e.g., bricklayer) and a kinship term (e.g., mother) could apply to the same person. The authors found that individuals exposed to counter-stereotypical images displayed faster reaction times and accuracy on judgments of incongruent word pairs than did those exposed to stereotypical images (Finnegan et al., 2015). In the present research, this image intervention will be applied to a different implicit measure – the DRM paradigm and the adapted gender role lists created by Lenton, Blair, and Hastie (2000). Additionally, this image intervention will also be applied to the inward measures of stereotypes by assessing its influence on math performance.

The current research seeks to address all of these gaps and combine the existing literature by exposing participants to either stereotypical or counter-stereotypical gender role images (used by Finnegan et al., 2015) and examining the effects on false stereotypical memories (using the DRM paradigm and the adapted gender lists used by Lenton et al., 2000) and stereotype threat. Based on the existing literature regarding implicit stereotypes and stereotype threat, we hypothesized that individuals who are exposed to counter-stereotypical gender role images (as opposed to those exposed to stereotypical images) will: a) recognize fewer false stereotypical memories (Hypothesis 1), and b) be less vulnerable to stereotype threat (Hypothesis 2).

STUDY 1

Method

Study 1 was intended to get an estimate of the current rate of false memories among an online sample of adults using the DRM paradigm. Gender stereotypical false memories were of specific interest in this experiment. Lenton, Blair, and Hastie's (2000) study was conceptually replicated in order to give an estimate of current implicit gender role associations in an online population (the original study utilized undergraduate students). Rate of indirect gender associations (i.e., implicit gender stereotypes) was the focus in this study.

Using Roediger and McDermott's (1995) DRM paradigm, direct associations were measured; in addition, Lenton, Blair, and Hastie's (2000) adapted gender list measured indirect associations. See Appendix A for word lists. We hypothesized that participants would have higher rates of false memories for gender role words that were consistent with the gender stereotypical theme that the participants were exposed to (male or female). Since social roles are often associated with particular traits, we also hypothesized that participants would falsely recognize stereotypic traits that correspond with the gender of the list they were exposed to. For example, participants exposed to a female list should falsely recognize roles and traits that are stereotypically feminine, such as hairdresser and caring, at a similar rate as female words that were actually presented in the list, while having a lower rate of recognition for roles and traits that are stereotypically masculine, such as engineer and strong.

Procedure

The methodology of Study 1 replicated that of Lenton, Blair, and Hastie’s (2000) study which measured implicit associations among participants. However, the current experiment was administered using Qualtrics (a survey software capable of collecting data from online populations) via MTurk, instead of a lab setting using undergraduate students. After obtaining informed consent, participants were randomly assigned to either a male or female stereotype list condition. The experiment then continued in four phases: exposure phase, recognition test, theme awareness, and demographic information.

Exposure phase. During the exposure phase, participants serially viewed a list of words, and were instructed to study each word for later recognition. Each word was presented for 2 seconds, with no break between word lists. The word lists were provided by the DRM paradigm (Roediger & McDermott, 1995), and also included the two adapted gender role word lists created by Lenton, Blair, and Hastie (2000) (Appendix A).

Participants viewed a total of 90 words; 6 categories containing 15 words each were presented without breaks between categories, appearing as one long list of words. Following Lenton et al. (2000), word lists were presented in the same order for all participants so as to avoid any potential confound concerning order: (1) *chair* list, (2) *fruit* list, (3) *stereotype* list (either male or female – randomly assigned), (4) *window* list, (5) *sleep* list, (6) *gender-neutral trait* list. After all words were presented, participants worked on a distractor task for 2 minutes, which consisted of answering basic mathematical problems (see Appendix D for problems). This distractor task simply prevented participants from rehearsing words.

Recognition test. After participants worked on the distractor task for 2 minutes, the recognition test was administered. Test words appeared on the screen, and participants were asked to decide on their recognition of each word by indicating if the word was *Definitely new, Probably new, Probably old,* or *Definitely old.* The recognition test contained 50 items consisting of 12 potential hits (i.e., words actually presented in the exposure phase; 2 from each list), 4 direct lures (i.e., words not presented in the exposure phase but are related to each of the non-gendered words; 1 from each direct list), 4 masculine role lures, 4 masculine trait lures, 4 feminine role lures, 4 feminine trait lures, and 18 (non-studied) filler words not closely associated with any of the words presented in the exposure phase (Appendix A). All participants received the same recognition test no matter which word list they received.

Theme awareness. After participants completed the recognition test, they were asked about their explicit awareness of word themes. Participants were instructed about what a word theme is, and then asked to list any word themes that they may have noticed or used in the recognition test phase. This allowed us to determine if participants used explicit schema information during the recognition test (e.g., stereotypically "female jobs"). Like the findings of Lenton et al. (2000), we expected that most participants would not report using any explicit schemas for the gender lists, which would support the role of implicit indirect associations.

Demographic information and BSRI. Participants were asked basic demographic information and completed the Bem Sex Role Inventory (BSRI) to measure individual schematic differences (i.e., self-endorsement of masculine and feminine characteristics; see Appendices B and C). Finally, participants were debriefed, thanked for their participation, and received instructions to collect their compensation through Amazon MTurk.

Design and Analytic Strategy

Results were analyzed using IBM SPSS version 25. We measured the ratings of false alarms (incorrectly identifying a new word as old) and hits (correctly identifying an old word as old) for each word list. Additionally, we measured both direct false alarms (i.e., words directly related to each other; from Roediger & McDermott's 1995 original lists) and indirect false alarms (i.e., words related through a third, mediating concept; from Lenton, Blair, & Hastie's 2000 adapted gender role list). For our main hypothesis, we were particularly interested in scores of indirect false alarms as a measure of implicit associations. Response options for each word ranged from 1 (*definitely new*) to 4 (*definitely old*). For all analyses, word lists were standardized by dividing the total score by the number of items to allow for comparison between word type categories because there were different numbers of items in the conditions (the reader is directed to Appendix A for a reminder of word list items).

We first examined correct and false recognition of non-gendered items to gain a baseline estimate of general false memory rates by conducting a 2 (gender of participant) x 2 (gender of presented word list) x 3 (word type: nonstudied critical lures, nonstudied filler words, studied category [non-gendered] words) mixed design ANOVA, with repeated measures on the last factor. To test our primary research question of Study 1, we then examined if recognition of stereotypically masculine and feminine roles and traits varied by gender of participant and gender of presented word lists by conducting a 2 (participant gender) x 2 (gender of presented word lists) x 2 (recognition test word type: role, trait) x 2 (recognition test word gender) mixed-design ANOVA, with repeated measures on the last 2 factors. Finally, we examined how implicit gender stereotypes (i.e., false recognition of stereotype-consistent words) were associated with participants' self-endorsement of gender stereotypical characteristics (masculine, feminine, and neutral) using the BSRI. We computed separate masculinity and femininity scores for each participant by using the average scores for masculine and feminine items (as detailed by Bem, 1974). We then conducted separate correlations for each of the gender word lists presented with masculine and feminine BSRI scores. For all analyses, descriptive statistics were examined to determine the direction of relationship in significant interactions.

We hypothesized that our results would support those found by Lenton, Blair, and Hastie (2000). Specifically, we hypothesized that exposure to either male or female lists would result in significantly higher false alarms to stereotype-consistent lures than stereotype-inconsistent lures (i.e., participants exposed to the female list condition would have more false alarms for female lures than those exposed to the male list, and participants exposed to the male list condition would have more false alarms for male lures than those exposed to the female list). Additionally, we hypothesized that participants' existing gender views would play a role in rates of false memories. Because gender schema theory posits that individuals with more traditional gender views should use existing gender schemas more when interpreting gendered information, we hypothesized that participants who endorse more traditional gender roles (as measured by the BSRI) would display higher rates of false stereotypical memories than those with less traditional gender roles.

Participants

We recruited participants through Amazon Mechanical Turk (MTurk), an online crowdsourcing website with a large and diverse sampling pool through which companies and social scientists can compensate participants to perform tasks, such as completing surveys (Buhrmester, Kwang, & Gosling, 2011). The task description simply asked participants to complete a memory test. Potential participants were required to be at least 18 years of age, live in the U.S., and have English as a first language. The study took approximately 20-30 minutes to complete, and participants were compensated 75 cents for their participation. A total of 161 MTurk users participated in our study. The data of

60 participants had to be excluded from analyses, however, for the following reasons: not meeting inclusion criteria (nonnative speaker), not completing the recognition test, noticing a gender theme in the word lists, and reporting specific distractions during the study. Data were analyzed for the remaining 101 participants who completed the study.

Nearly half of our participants (46%) identified as female. A majority identified as heterosexual (87%), were either married (48%) or single (32%), and held at least a bachelor's degree (58%). Participants ranged in age from 20 to 70 years old ($M = 36.27$).

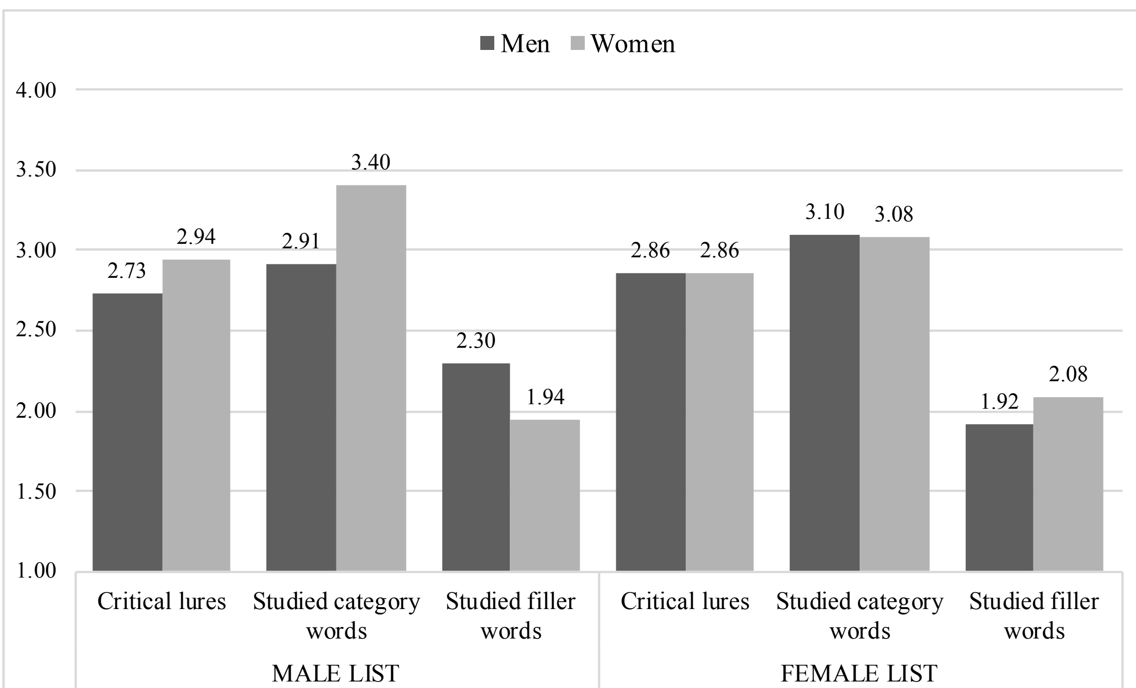
Results

Non-Gendered Word List Items (i.e., Direct Associations)

Across all categories, the studied category words (e.g., berry, juice, nap, blanket, etc.) were correctly recognized as old ($M = 3.21$). The non-studied critical lure words (e.g., sleep) were more likely to be falsely recognized as old than new ($M = 2.84$), and the non-studied filler words were less likely to be falsely recognized ($M = 2.08$). This difference in recognition by type of word category was significant, $F(2, 180) = 65.66$, $p < .001$, $\eta^2 = .42$. These findings replicate those of Deese, Roediger, and McDermott (1995) and demonstrate participants' false memory and implicit associations of non-gendered words at a rate above neutral (a rating of 2.84 on a scale ranging from 1 to 4).

The interactions between type of word category and list gender, as well as the type of word category and participant gender were both non-significant. In other words, mean recognition scores for the word categories did not significantly vary by participant gender or by gender of the presented word lists (p's > .26). However, we did find a significant three-way interaction among word category, participant gender, and list gender, $F(2, 180) = 3.82$, $p = .024$, $\eta^2 = .04$ (see Figure 1). Specifically, when presented with the female list, men and women recognized the categories at about the same rate. When presented with the male list, however, women were less likely to recognize the non-studied filler words, but more of the studied category words than men. Additionally, women were more likely to falsely recognize the critical lures when presented with the male list. These findings suggest possible gender differences in correct and false recognition of even non-gendered word list items, but only when presented with particular gendered stimuli.

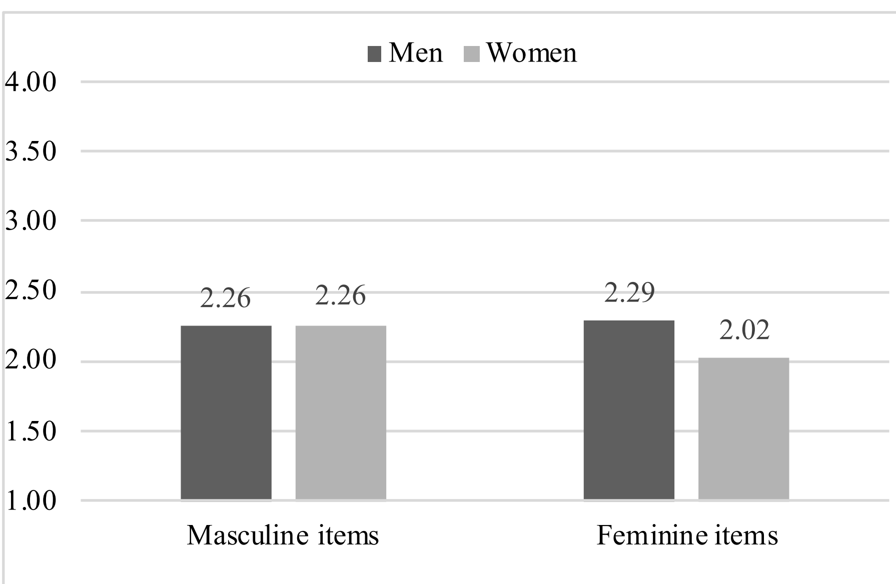
Figure 1. Correct and False Recognition of Non-Gendered Words by Participant Gender and List Gender.



Gendered Word List Items (i.e., Indirect Associations / Implicit Gender Stereotypes)

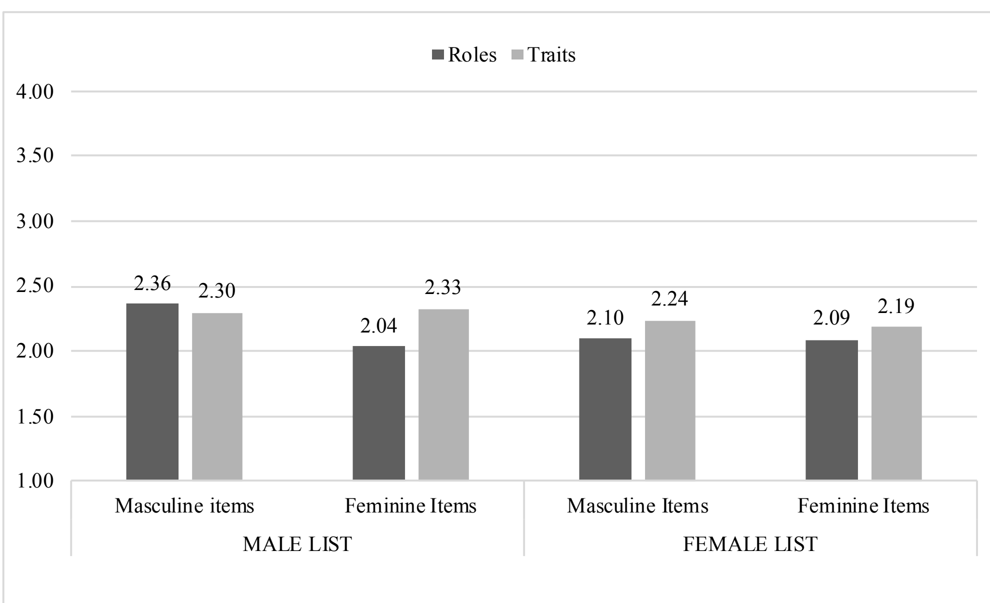
When we examined our main research question, we discovered several significant results. First, overall, masculine items were more likely to be falsely recognized (M = 2.26) than feminine items (M = 2.16), $F(1, 93) = 5.95$, $p = .017$, $\eta^2 = .06$. Second, traits were more likely to be falsely recognized (M = 2.27) than roles (M = 2.16), $F(1, 93) = 4.93$, $p = .029$, $\eta^2 = .05$. Third, there was a significant item gender x participant gender interaction, $F(1, 93) = 12.75$, $p = .001$, $\eta^2 = .12$ (see Figure 2). Specifically, men and women were equally as likely to falsely recognize masculine items (M's of both = 2.26), and men were about as likely to falsely recognize masculine items as feminine items (M = 2.29). However, for women, there was a significant item gender effect, $F(1,43) = 18.96$, $p < .001$, $\eta^2 = .31$, demonstrating that women were less likely to falsely recognize feminine items (M = 2.02) than were men (M = 2.29).

Figure 2. Recognition of Masculine and Feminine Words by Participant Gender.



Finally, there was a significant three-way interaction among gender item, item type, and list gender, $F(1, 93) = 4.28$, $p = .041$, $\eta^2 = .04$ (Figure 3). When presented with a female list, masculine and feminine roles were both recognized as "new" at approximately the same rate (M's = 2.10 and 2.09, respectively), but masculine traits were falsely recognized at a higher rate ($M = 2.24$) than feminine traits ($M = 2.19$). When presented with a male list, however, masculine roles were much more likely to be falsely recognized ($M = 2.36$) than feminine roles ($M = 2.04$), but the rate of false recognition of masculine and feminine traits was approximately equal (M's = 2.30 and 2.33, respectively). Note that the participant gender was not part of this pattern as indicated by a nonsignificant quadruple interaction, implying that both men and women are showing similar biases in recognition. Because of this triple interaction, male and female lists were examined separately.

Figure 3. Recognition of Masculine and Feminine Words by Participant Gender and Word Type.



In these follow-up ANOVAs, we found that when the male list was presented to participants, masculine items were significantly more likely to be falsely recognized (M = 2.33) than feminine items (M = 2.18), $F(1, 58) = 7.11$, $p = .010$, $\eta^2 = .11$. We also discovered a significant item gender x item type interaction, $F(1, 58) = 7.97$, $p = .007$, $\eta^2 = .12$. Feminine and masculine traits showed similar levels of false recognition (M's = 2.33 and 2.30, respectively) but the pattern changed for the roles; masculine roles were more likely to be falsely recognized (M = 2.36) than feminine roles (M = 2.04), $F(1, 60) = 13.59$, $p < .001$, $\eta^2 = .19$. When the female list was presented, however, the item gender x item type interaction was no longer significant ($p > .10$), so we did not analyze this interaction further for the female list. It is also worth noting that since the baseline false recognition was 2.1 for non-gendered items, it can be assumed that false recognition of gendered items was low in the female list.

Existing Gender Views

In general, participants self-endorsed fairly traditional gender views as evidenced by BSRI scores. On the 7-point scale ranging from 1 (*never or almost never* true) to 7 (*always or almost always* true), the average femininity score was 4.54 (range: 2.10 to 6.50) and the average masculinity score was 4.40 (range: 2.40 to 6.60). These scores varied by gender: men demonstrated significantly higher masculine scores than women (men's M = 4.67, women's M = 4.08), $t(98) = 3.28$, $p = .001$, and women demonstrated significantly higher feminine scores than men (men's M = 4.26, women's M = 4.87), $t(98) = -3.83$, $p < .001$. Contrary to our hypothesis, however, masculine and feminine

BSRI scores were not significantly correlated with participants' rate of false memories on any of the recognition list types (all *p*'s > .13).

Study 1 Discussion

Our results from Study 1 demonstrated several important findings. First, the high rate of false alarms of non-gendered items replicated those of the original DRM paradigm (Roediger & McDermott, 1995) among an online sample, decades after the original study. Although not novel, these findings support the cognitive paradigm of false memory and implicit bias and demonstrate their pervasiveness in everyday life.

Second, we discovered some important interactions and gender differences among participant performance. Overall, masculine items were more likely to be falsely recognized than feminine items. And although men and women were about equally as likely to falsely recognize masculine items, women were less likely to falsely recognize feminine items than were men. This pattern suggests that both men and women are affected by male stereotypes, but when it comes to female stereotypes men are more affected than women. Stereotypically feminine words may be more narrowly associated with women, but masculine items may be associated more generously with either gender.

We also found that roles and traits were differentially recognized depending on the gender of the list presented to participants. Interestingly, in the male list, there was a stronger false memory effect and an interaction. In the male list, both masculine traits and roles were falsely recognized, but only feminine traits (not roles) showed false recognition. In the female list, overall, there was little false recognition with both feminine and masculine traits showing slightly higher false recognition than roles. It is also important to note that participant gender did not significantly influence this interaction. These findings across the two list types suggest that presenting participants with masculine cues may promote a sharpened awareness of societal roles, as demonstrated by the higher rates of false stereotypical memories for masculine roles compared to feminine roles when presented with stereotypically masculine word lists and a lack of such an interaction when presented with a stereotypically feminine word list. Additionally, they may hint at a narrower schema of masculine roles compared to feminine roles but a broader schema of masculine traits, regardless of the gender of the perceiver.

Our lack of significant associations between existing gender views and false stereotypic memories suggests that participants with more traditional gender views may not have used their existing gender schema when processing the word lists any more than participants with less traditional gender views. These findings may be explained, however, by the fact that very few participants ($n = 3$) were explicitly aware of a gendered schema in the word lists and these participants were excluded from analyses. Thus, although gender schema theory posits that gender-based schematic processing should result in more gender-consistent associations, this type of schematic processing may not apply to associations made out of conscious awareness. In other words, gender schema theory may not be able to explain false stereotypic memories created by the use of an implicit schema, at least in terms of the DRM paradigm.

STUDY 2

Method

Study 1 provided data regarding the rates of false memories (specifically, the number of indirect implicit gender stereotypes) using an adapted version of the DRM paradigm. In order to learn more about factors that may influence these false memories, Study 2 examined the effects of a simple image priming intervention on the activation of these indirect implicit gender stereotypes using the same word lists as Study 1. Specifically, the effects of viewing stereotypical vs counter-stereotypical role images (proposed by Finnegan et al., 2015) on the occurrence of false memories using the DRM paradigm and Lenton et al.'s (2000) adapted gender role and trait lists were examined. Indirect gender role and trait associations were measured after participants were exposed to either stereotypical or counter-stereotypical role images.

Procedure

After obtaining informed consent, participants were randomly assigned to either a stereotypical or counter-stereotypical image condition. The study then continued in five phases: image viewing, exposure phase, recognition test, theme awareness, and demographic information.

Image viewing. Similar to Finnegan et al. (2015), participants serially viewed a series of 24 images, reflecting either stereotypical (e.g., female make-up artist) or counter-stereotypical (e.g., male make-up artist) gender roles (see Appendix E for images). These were the same images already used and established by Finnegan et al. (2015). While each image was displayed on screen, 2 brief sentences were presented describing the character in the image (e.g., "This is Paul. He is a make-up artist"). In order to focus participants' attention on each image, they were asked to answer four questions (two open-ended, two with response options) relating to each character's probable salary, job satisfaction, leisure activities, and lifestyle.

Participants were then randomly assigned to either a female or male stereotype list condition, and the remainder of the experiment continued in the same order and used the same methodology as Study 1: exposure phase, recognition test, theme awareness, and demographic information. Participants were then debriefed, thanked for their participation, and received instructions to collect their compensation through Amazon MTurk.

Design and Analytic Strategy

Similar to Study 1, the rate of false alarms for each list was measured. The primary dependent variables of interest were the false alarm scores from the gender role and trait lists. Results were analyzed using a 2 (image type: stereotypical, non-stereotypical) x 2 (list gender: masculine, feminine) x 2 (participant gender) x 2 (item type: role, trait) x 2 (item gender: masculine, feminine) mixed-design ANOVA, with repeated measures on the last two factors. We hypothesized that participants exposed to counter-stereotypical role images would recognize fewer false gender stereotypical memories than participants exposed to stereotypical images. The image manipulation was not expected to influence the rate of false alarms to direct lures (e.g., bed and sleep).

Participants

Interested MTurk participants were asked to complete a memory test and indicate occupational opinions (as they were asked to answer brief questions on each image regarding personal happiness, income, etc. as a way to focus participants' attention on each image). After reading the study description and requirements (at least 18 years of age, live in the U.S., and have English as a first language), participants were redirected to Qualtrics to complete the study. The study took approximately 45 minutes to complete and participants were compensated 75 cents for their participation. A total of 425 participants agreed to participate. Unfortunately, however, the data of 277 of these participants had to be excluded from analyses for the following reasons: not meeting the inclusionary criteria (nonnative speaker), noticing a gender theme in the word lists, spending far too long to complete the study accurately (9+ hours), not completing the study through the recognition portion (which provided the main dependent variable), and for responding in a manner that resembled "bots".

During the time of our data collection, many social scientists that used the MTurk platform for their research reported a decrease in quality responses; upon further investigation, researchers found the culprit to be bots - "automated programs mimicking human behavior" (Dreyfuss, 2018). To locate bots in our study, we examined the open-ended question responses and GPS locations (the main complaints reported by social scientists as evidence of bot responses; Dreyfuss, 2018); participants with nonsensical responses and/or repeat locations were designated as bots and excluded from all analyses. Some examples of nonsensical responding involved repeated responses (often in all capital letters) to many or all open-ended questions about how the individuals in the priming images likely spend their free time that did not make sense for these questions (e.g., "IT'S GOOG LIFE", "YES", "Really my days are exhausting and did not stop working") and/or did not make sense grammatically and suggested that the respondent did not have English as a first language (e.g., "12 havays", "HE FULLY ENJOYING HIS JOB", "JOIN WITH HIS FAMILY"). We saw this as a rather conservative approach to detecting bots as we did not exclude responses that were not grammatically perfect (i.e., we left room for human error) and we did not exclude respondents that provided blank answers to open-ended questions. After excluding the data of participants that did not finish the recognition test first ($n = 250$), we discovered at least 30 respondents that we designated as bots.

In total, the data from 148 participants were used in analyses. A majority of our remaining participants were female (58%), Caucasian (78%), heterosexual (89%), either married (53%) or single (30%), and held at least a bachelor's degree (69%). Image condition and gender of list were roughly equally presented to participants: 67 participants viewed stereotypical images (72 viewed counter-stereotypical) and 64 participants viewed the male word list (64 viewed female).

Results

Non-Gendered Word List Items (i.e., Direct Associations)

When we examined the false recognition of non-gendered items, there was a significant item type x list gender interaction, $F(2, 240) = 6.85$, $\eta^2 = .05$. On studied words, the performance did not significantly vary as a function of the image type or the gender of the list that was presented; all four groups had an average of 3.13 on studied items (*F*'s < 1). On non-studied items, the non-stereotypical image group had more false recognition (M = 2.23 vs. 1.90). On the direct lures, stereotypical images led to more false memory on the masculine list than the feminine list (M = 3.11 and 2.66, respectively), $F(1, 66) = 5.50$, $p = .022$, $\eta^2 = .08$. However, with non-stereotypical images, there were no significant list gender differences (M = 3.14 vs. 2.97), $F(1, 73) = 1.36$. On non-studied items, the counter-stereotypical image group had higher false recognition (M = 2.23 vs. 1.90), $F(1,130) = 5.98$, $p = .016$, $\eta^2 = .04$. When the images activated unfamiliar, non-stereotypical careers, participants seemed confused and produced more false memory even for non-gendered items.

Gendered Word List Items (i.e., Indirect Associations)

When we examined our main research question, we found that participant gender did not produce significant differences on false memory by itself or in an interaction with any other variable; thus, we collapsed the results across both men and women. Image type also did not have a main effect or an interaction with other variables, contrary to our hypothesis. The only significant effects were an interaction between item gender and list gender, $F(1, 131) = 5.15$, $p < .05$, $\eta^2 = .04$ (Figure 4) and an interaction between item gender and item type, $F(1, 131) = 27.69$, $p < .001$, $\eta^2 = .18$ (Figure 5). The list gender and item gender interaction indicated that when the list had male characteristics, there was an increase in false recognition of masculine items (both roles and traits collapsed), $F(1,63) = 29.59$, $p < .001$, $\eta^2 = .32$. For the female list, masculine and feminine items had similarly low levels of false recognition ($F < 1$). The item gender and item type interaction showed that, for roles, there is a difference by item gender, $F(1,142) = 32.07$, $p < .001$, $\eta^2 = .18$. Male roles were falsely recognized more than female roles ($M = 2.24$ vs. 1.95). For traits, however, there was no significant difference by gender ($F < 1$).

Figure 4. Recognition of Masculine and Feminine Words by List Gender.

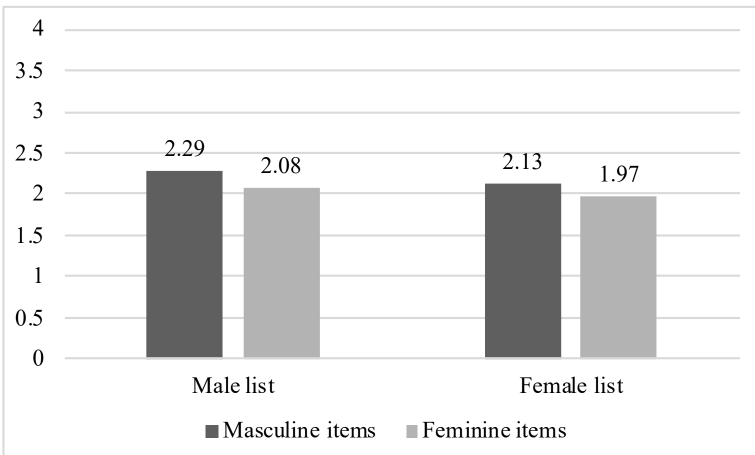
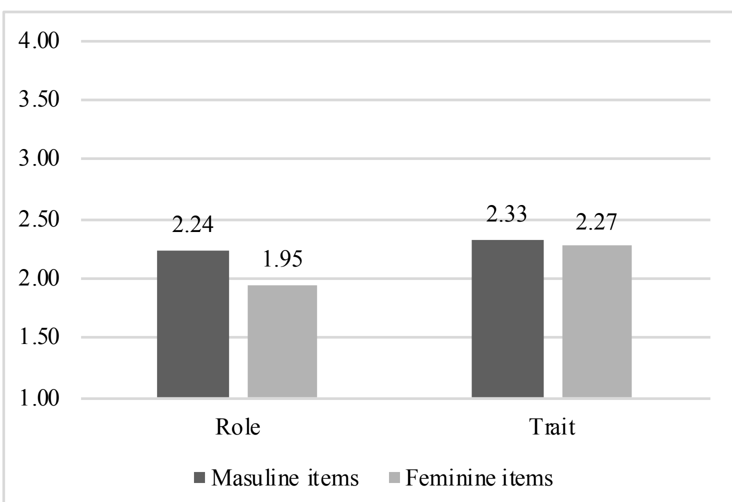


Figure 5. Recognition of Masculine and Feminine Words by Word Type.



Existing Gender Views

Participants again self-endorsed relatively traditional gender views: the average masculine score was 4.32 and the average feminine score was 4.52. Unlike in study 1, however, masculine and feminine scores were not significantly different by gender (i.e., men and women both self-endorsed masculine and feminine characteristics at roughly the same rate; p's > 10). Masculine and feminine BSRI scores were again not significantly correlated with performance on any of the word lists (p's > .08). The type of images presented to participants did not significantly influence BSRI scores (p's > .40).

Study 2 Discussion

Our data did not show any effects of the image manipulation on measures of implicit gender stereotypes. We also did not find that existing gender views influenced any of the false memory outcomes. Although our hypotheses were not supported, we did discover some interesting results that hint at some of the nuances involved in implicit gender stereotypes. For example, false recognition of gendered information may vary depending on the type of information presented (e.g., role or trait). Additionally, our results suggest that the gender of the perceiver may not influence implicit associations (i.e., men and women may be just as likely to experience biased perceptions toward either gender).

STUDY 3

Method

Study 2 examined the impact of exposure to stereotypical vs. counter-stereotypical images on the outward expression of gender role stereotypes. The goal of Study 3 was to examine the impact of exposure to such images on the more inward expression of gender stereotypes by measuring stereotype threat.

Participants viewed either stereotypical or counter-stereotypical images of gender roles (the same images used in Study 2), and completed a stereotype threat task, consisting of practice ACT mathematical questions (see Appendix F for questions). We hypothesized that participants exposed to counter-stereotypical images would demonstrate better mathematical performance (and thus be less likely to experience stereotype threat in this study). Additionally, the role of image exposure on feelings of math anxiety was explored (as measured by the Math Anxiety survey).

Procedure

After obtaining informed consent, participants were randomly assigned to either a stereotypical or counter-stereotypical image condition. As in Study 2, participants then serially viewed either stereotypical or counter-stereotypical images and answered the same four questions regarding each character's probable job satisfaction, salary, leisure activities, and lifestyle. Participants then completed the series of math questions, followed by a demographic questionnaire and the Math Anxiety survey (Appendix G). Finally, participants were debriefed, thanked for their participation, and received instructions for collecting their compensation through MTurk.

Participants

Interested participants were asked to indicate occupational opinions, complete a memory test, answer some math questions, and complete some brief surveys. After reading the study description and requirements (at least 18 years of age, live in the U.S., have English as a first language, and self-identify as female), participants were redirected to Qualtrics to complete the study. The study took approximately 60 minutes to complete and participants were compensated $1.00 for their participation. A total of 280 participants agreed to participate. A large number of responses, however, again had to be excluded from analyses for the following reasons: not meeting inclusion criteria (nonnative speaker), nonsensical responding, spending too long or too short of a time to reliably complete the study (< 10 minutes or > 2.5 hours), and for not answering any questions. For reasons not entirely known to the authors, nonsensical responding and responses in nonnative languages observed in the open-ended questions were more common in this study than in Study 2, which we attribute to the sharp rise in bot responding seen during the relatively short time period of data collection (Dreyfuss, 2018).

Regardless of the exact reason(s) for the high rate of nonsensical responses, data were available for the remaining 80 participants who demonstrated genuine and logical responding throughout the study. Of these participants, 25 identified as women and 22 as men; 33 did not complete the demographic survey at the end, leaving a total of only 47 participants in the sample. Remaining participants ranged in age from 21 to 61 years (M =

34.67). Most participants identified as heterosexual (91%), were married (60%), and held at least a bachelor's degree (67%).

Design and Analytic Strategy

Due to the high number of excluded responses, low cell sizes and low power became important issues to consider in analyses. Thus, our sample was underpowered and not likely to detect any meaningful differences in math performance in women based on the type of images presented and/or existing gender views via ANOVA. Therefore, instead, Pearson correlations were conducted to determine which factors were associated with math anxiety and math performance, irrespective of gender.

We first conducted a chi-square test to examine the potential impact of image type on the likelihood of completing the math portion of the study to investigate the potential confounding issue of one group self-selecting to answer the math questions based on image type (e.g., if participants who were presented with stereotypical images were more likely to discontinue the study than participants presented with non-stereotypical images).

We then conducted correlations between relevant variables; specifically, type of images presented, math performance (math anxiety, number of math problems answered, number of math problems answered correctly), math experience (how often the participant uses math in their career and daily life, feelings toward math, feelings about their personal math ability, the number of math courses completed) and demographic variables (participant gender, education). The math anxiety scale was coded and calculated so that higher scores indicated higher math confidence (i.e., lower math anxiety). Table 1 below shows which measure each variable was taken from and how each was calculated, and Appendices B, G, and H show the exact questions and response options used.

Results

Chi-Square

The results of the chi-square test indicated no significant differences between type of image presented (stereotypical vs. counter-stereotypical) on the likelihood of completing the math problems ($p < .20$).

ANOVA

Due to small sample sizes in each group, a 2 (image type) x 2 (gender) ANOVA did not result in significant differences in math performance, $F(1, 42) = 2.51$, $p > .10$, $\eta_p^2 = .06$. Upon examining confidence intervals (CI's), however, our results hint at an interesting pattern (see Table 1 for mean scores). Despite having relatively similar levels of education, and experience and comfort with math, the men who saw counter-stereotypical images tended to do more poorly on math questions (*M* number of correct answers = 6.79 [95% CI 5.96, 7.62]) than men in the stereotypical image group (*M* = 9.38 [95% CI 6.01, 12.75]) and the two image groups of women (stereotypical *M* = 8.70 [95% CI 7.15, 10.25]; counter-stereotypical *M* = 9.07 [95% CI 7.32, 10.82]). These results suggest that men exposed to counter-stereotypical images performed worse than men exposed to stereotypical images *and* women exposed to both types of images.

Table 1. Mean *(sd)* Scores of Math Variables by Gender and Image Condition.

		Stereotypical	Counter-stereotypical
Men	N:	8	13
	# correct:	9.38 *(4.87)*	6.77 *(1.64)*
	Math comfort:	3.03 *(.95)*	2.79 *(.69)*
	Use math career:	2.13 *(.64)*	2.38 *(1.21)*
	Use math daily:	2.25 *(.46)*	2.46 *(.88)*
	Math feelings:	3.13 *(.99)*	3.38 *(1.39)*
	Math ability:	2.88 *(1.13)*	4.15 *(2.08)*
	# math courses:	6.75 *(3.54)*	6.31 *(3.33)*
Women	N:	10	14
	# correct:	8.70 *(2.50)*	9.07 *(3.34)*
	Math comfort:	2.54 *(.86)*	2.99 *(.71)*
	Use math career:	2.20 *(.79)*	2.29 *(.99)*
	Use math daily:	2.20 *(.92)*	2.79 *(.89)*
	Math feelings:	2.60 *(1.43)*	3.36 *(1.51)*
	Math ability:	3.10 *(1.91)*	3.14 *(1.56)*
	# math courses:	6.30 *(3.06)*	6.29 *(2.76)*

Note. Math comfort scores are based on mean scores from all items in the Math Anxiety scale, ranging from 1 to 5 and coded so that a higher score = more comfort (lower anxiety).

Correlations

See Table 2 for correlations of relevant variables and Table 3 for how these variables were calculated.

Table 2. Pearson Correlations of Demographic, Math Performance, and Attitude Variables.

	1	2	3	4	5	6	7	8	9
1. Math comfort									
2. # Correct answers	.44**								
3. Career math use	.34*	-0.10							
4. Daily math use	.34*	-0.15	.69***						
5. Feelings toward math	0.76***	0.23	.63***	.59***					
6. Feelings of personal math ability	.68***	0.10	.61***	.48**	.80***				
7. Education	.38*	-0.02	.31*	.32*	.40**	.30*			
8. # Math courses	.34*	0.23	0.11	0.18	0.28	0.23	0.02		
9. Image type	0.11	0.04	0.10	0.24	0.21	0.18	0.06	-0.06	

*p < .05. **p < .01. ***p < .001.

Table 3. Variables Used in Correlations.

Variable	Measure	Calculation
Math comfort	Math Anxiety Survey	Higher score = higher math comfort
Number of correct answers	ACT practice questions	Higher score = more correct answers
Career math use	Demographic question	Higher score = more frequent use
Daily math use	Demographic question	Higher score = more frequent use
Feelings toward math	Demographic question	Higher score = more positive
Feelings of personal math ability	Demographic question	Higher score = more competent
Education	Demographic question	Higher score = higher education
Number of math courses	Demographic question	Higher score = more courses

Gender and Stereotypical Images

Contrary to our hypotheses, neither participant gender nor image type were significantly correlated with any of our relevant variables. Both genders had similar patterns related to math anxiety and performance. Additionally, the number of math problems answered correctly was not significantly correlated with any other variable.

Math Comfort

Our results showed some significant correlations between demographic characteristics and math comfort across both genders. First, math comfort was correlated with math performance, so we explored math comfort's correlations with other variables.

Math comfort was significantly correlated with many variables in directions that would be expected (see Table 2). Specifically, participants with lower math anxiety (i.e., higher math comfort) tended take more math classes, use math more often in their daily lives and careers, have more positive feelings toward math in general and toward their own personal math abilities, and had completed a higher level of education. It is interesting to note that two simple questions – one about participants' feelings toward math in general and one about their perceptions of their own math skills – correlated very strongly with the math anxiety survey scores.

Study 3 Discussion

Despite the low power we experienced in Study 3, we were able to find some interesting results. First, although participant gender and image type were not significantly correlated with any relevant variables and did not support our hypotheses, we view it as a positive finding that gender was not associated with education, math performance, math use, and feelings toward math. When women who are comfortable with math self-select into the study, they do not demonstrate statistical differences in terms of math experience, attitudes, and performance compared to men.

Also, the fact that we discovered many significant correlations with math comfort despite our low cell sizes suggests that math anxiety holds important implications for math experiences and performance. Although we cannot determine the directionality of the math experience and math comfort relationship, these associations highlight the importance of learning more about techniques to reduce math anxiety.

GENERAL DISCUSSION AND LIMITATIONS

Overall, we discovered in the first two studies that false memory could be demonstrated with lures that were directly related to themes, but the false memory performance was considerably weaker for indirect lures (gender roles and traits), providing only a partial replication of Lenton et al. (2000). The false memory for gender-related items tended to be stronger when the lists biased the individuals to consider male concepts. Another lack of conceptual replication was the effectiveness of counter-stereotypical images to reduce implicit bias. Our data did not replicate Finnegan et al.'s (2015) results with a different paradigm; there may be several reasons for this. First, the counter-stereotypical image technique may not work with a false memory measure as that measure of bias may be too week to yield any effects of an intervention. It is also worth noting that the DRM paradigm uses a distractor task between the exposure phase and the recognition phase to keep participants from rehearsing words, but Finnegan et al. (2015) did not use such a distractor task in their procedure. It is possible, then, that including a distraction before an implicit measure produces smaller (but perhaps more reliable) effects.

Additionally, it is possible that such a brief image intervention is not a strong enough manipulation to provide large effect sizes in the reduction of implicit stereotypes. Although some studies have shown that a single intervention session can reduce implicit stereotypes, the effect sizes found are generally small and do not completely illuminate stereotypes (Blair, 2002; Gregg, Seibt, & Banaji, 2006; Lenton, Sedikides, & Bruder, 2009). Thus, it is possible that reducing automatic stereotypes involves a learning process that requires multiple counter-stereotypical exposure experiences over time (Lenton et al., 2009). Finally, another large limitation in our study was the lack of power given the number of ineligible participants in Studies 2 and 3.

IMPLICATIONS AND FUTURE DIRECTIONS

Our study examined the replicability of previous studies examining stereotypes and provided some interesting information on indirect associations. It could be expanded upon, however, by continuing to investigate which and how variables work to influence implicit bias. For example, future studies could examine if our lack of statistically significant relationships were actually caused by low power, or if such an image manipulation was too weak or unsuitable for indirect associations. More broadly, because our use of automatic cognitive processes (i.e., heuristics) may be more likely to be used in ambiguous situations or times of cognitive overload (Tversky & Kahneman, 1974), future studies could manipulate the type of situation and level of cognitive strain to investigate the situations that are most likely to activate implicit stereotypes. It would also be worthwhile to investigate and disentangle how identification in multiple groups (e.g., women of color) works to influence the outward and inward expression of implicit stereotypes, as individuals of multiple minority status are disproportionately affected by bias; for example, although White women earned 77% of their male counterparts in 2017, Black women earned only 61% and Hispanic women earned 53% of what White men earned (Hegewisch, 2018). Finally, because implicit stereotypes appear to be learned from repeated exposure (Rinehart, 1963), it is also important to study when and how these biases develop in order to learn how to best reduce the development of implicit bias from a young age. As long as the application of implicit bias (e.g., discrimination) occurs toward groups of disadvantaged individuals, this subject is important to continue to study in the work toward equality.

References

Akrami, N. & Ekehammar, B. (2005). The association between implicit and explicit prejudice: The moderating role of motivation to control prejudiced reactions. *Scandinavian Journal of Psychology, 46,* 361-366.

Bem, S. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88,* 354-364.

Blair, I. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242-261.

Blair, I., Ma, J.,& Lenton, A. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81,* 828-841.

Brambilla, M., Ravenna, M., & Hewstone, M. (2012). Changing stereotype content through mental imagery: Imagining intergroup contact promotes stereotype change. *Group Processes and Intergroup Relations, 15,* 305-315.

Buhrmester, M., Kwang, T., & Goslin, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science, 6,* 3-5.

Campbell, S., & Collaer, M. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychology of Women Quarterly, 33,* 437-444.

Dreyfuss, E. (2018, August 17). *A bot panic hits Amazon's Mechanical Turk.* Retrieved from https://www.wired.com/story/amazon-mechanical-turk-bot-panic/

Finnegan, E., Oakhill, J., & Garnham, A. (2015). Counter-stereotypical pictures as a strategy for overcoming spontaneous gender stereotypes. *Frontiers in Psychology, 6,* 1-15.

Finnigan, K. & Corker, K. (2016). Do performance avoidance goals moderate the effects of different types of stereotype threat on women's math performance? *Journal of Research in Personality, 63,* 36-43.

Fogliati, V. & Bussey, K. (2013). Stereotype threat reduces motivation to improve: Effects of stereotype threat and feedback on women's intentions to improve mathematical ability. *Psychology of Women Quarterly, 37,* 310-324.

Franceschini, G., Galli, S., Chiesi, F., & Primi, C. (2014). Implicit gender-math stereotype and women's susceptibility to stereotype threat and stereotype lift. *Learning and Individual Differences, 32,* 273-277.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology, 24,* 645-662.

Gregg, A. P., Seibt, B., & Banaji, M. A. (2006). Easier said than done: Asymmetry in the malleability of automatic preferences. *Journal of Personality and Social Psychology, 90,* 1-20.

Greenwald, A., McGhee, D., & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

von Hippel, C., Sekaquaptewa, D., & McFarlane, M. (2015). Stereotype threat among women in finance: Negative effects on identity, workplace well-being, and recruiting. *Psychology of Women Quarterly, 39,* 405-414.

Hegewisch, A. (2018). The gender wage gap: 2017; earnings differences by gender, race, and ethnicity. Retrieved from https://iwpr.org/publications/gender-wage-gap-2017/

Hudak, M. (1993). Gender schema theory revisited: Men's stereotypes of American women. *Sex Roles, 28,* 279-293.

Kawakami, K., Dovidio, J., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78,* 871-888.

Lai, C., Marini, M., Lehr, S., Cerruti, C., Shin, J., Joy-Gaba, J., … Nosek, B. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General, 143*(4), 1765-1785.

Lenton, A., Blair, I., & Hastie, R. (2000). Illusions of gender: Stereotypes evoke false memories. *Journal of Experimental Social Psychology, 37,* 3-14.

Lenton, A. P., Sedikides, C., & Bruder, M. (2009). A latent semantic analysis of gender stereotype-consistency and narrowness in American English. *Sex Roles, 60,* 269-278.

Liben, L., Bigler, R., & Krogh, H. (2002). Language at work: Children's gendered interpretations of occupational titles. *Child Development, 73*(3), 810-828.

Luong, K. & Knobloch-Westerwick (2017). Can the media help women be better at math? Stereotype threat, selective exposure, media effects, and women's math performance. *Human Communication Research, 43,* 193-213.

Macrae, C., & Bodenhausen, G. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93-120.

McGlone, M., Aronson, J., & Kobrynowicz, D. (2006). Stereotype threat and the gender gap in political knowledge. *Psychology of Women Quarterly, 30,* 392-398.

Neuburger, S., Ruthsathz, V., Jansen, P., & Quaiser-Pohl, C. (2015). Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. *Learning and Individual Differences, 37,* 169-175.

Nosek, B., Greenwald, A., & Banaji, M. (2005). Understanding and using the implicit association test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31,* 166-180.

Rinehart, J. W. (1963). The meaning of stereotypes. *Theory Into Practice, 2,* 136-143.

Roediger, H., & McDermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803-814.

Shaffer, E., Marx, D., & Prislin, R. (2013). Mind the gap: Framing of women's success and representation in STEM affects women's math performance under threat. *Sex Roles, 68,* 454-463.

Smith, J., Brown, E., Thoman, D., & Deemer, E. (2015). Losing its expected communal value: How stereotype threat undermines women's identity as research scientists. *Social Psychology of Education, 18,* 443-466.

Spencer, S., Logel, C., & Davies, P. (2016). Stereotype threat. *Annual Review of Psychology, 67,* 415-437.

Spencer, S., Steele, C., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4-28.

Steele, C. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Attitudes and Social Cognition, 69,* 797-811.

Todd, A., Bodenhausen, G., Richeson, J., & Galinsky, A. (2011). Perspective taking combat automatic expressions of racial bias. *Journal of Personality and Social Psychology, 100,* 1027-1042.

Tomasetto, C. & Appoloni S. (2013). A lesson not to be learned? Understanding stereotype threat does not protect women from stereotype threat. *Social Psychology of Education, 16,* 199-213.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124-1131.

U.S. Department of Commerce (2011). *Women in STEM: A gender gap to innovation.* Washington, D.C.

U.S. Department of Education (2012). *Gender equity in education: A data snapshot.* Washington, D.C.

U.S. Department of Labor, Women's Bureau (2014). *Data & Statistics.* Washington, D.C.

Workplace Gender Equality Agency (2016). *International gender equality statistics.* Sydney, Australia.

Appendix A

Word Lists Used in Experiments 1-3

Chair
- table
- sit
- legs
- seat
- couch
- desk
- recliner
- sofa
- wood
- cushion
- swivel
- stool
- sitting
- rocking
- bench

Fruit
- apple
- vegetable
- orange
- kiwi
- citrus
- ripe
- pear
- banana
- berry
- cherry
- basket
- juice
- salad
- bowl
- cocktail

Window
- door
- glass
- pane
- shade
- ledge
- sill
- house
- open
- curtain
- frame
- view
- breeze
- sash
- screen
- shutter

Sleep
- bed
- rest
- awake
- tired
- dream
- wake
- snooze
- blanket
- doze
- slumber
- snore
- nap
- peace
- yawn
- drowsy

Female list
- secretary
- nurse
- housekeeper
- nanny
- assistant
- homemaker
- receptionist
- dietitian
- teacher
- therapist
- babysitter
- typist
- servant
- cashier
- model

Male list
- president
- detective
- coach
- mechanic
- soldier
- executive
- sheriff
- doctor
- farmer
- lawyer
- athlete
- rancher
- firefighter
- judge
- guard

Feminine trait lures
- warm
- caring
- delicate
- sensitive

Masculine trait lures
- active
- wise
- independent
- strong

Gender-neutral traits
- funny
- earnest
- articulate
- talented
- honest
- ethical
- adaptable
- happy
- candid
- punctual
- normal
- truthful
- enthusiastic
- sincere
- creative

Feminine role lures
- hairdresser*
- librarian*
- dancer*
- cheerleader

Masculine role lures
- engineer
- carpenter*
- architect*
- minister

**Note.* These role items were changed in the recognition test portion of Study 2 in order to avoid the potential confound of asking about participants' recognition of roles that were presented in the image priming. These items were changed to dressmaker (*hairdresser*), social worker (*librarian*), dental hygienist (*dancer*), bus driver (*carpenter*), and plumber (*architect*).

Appendix B

Demographic Questions Used in Experiments 1-3

- What is your age in years?
  - (Open response)
- What gender do you identify as?
  - Male
  - Female
  - Other: ______________
- Which of the following best describes your sexual orientation?
  - Heterosexual (straight)
  - Gay
  - Lesbian
  - Bisexual
  - Other: _______________
- What is your relationship status?
  - Single, never married
  - Dating
  - Married
  - Widowed
  - Divorced/separated
  - Other: ________________
- What is your occupation?
  - (Open response)
- What is your individual annual income?
  - (Open response)
- What is your highest level of education completed?
  - No schooling completed
  - Some high school, no diploma
  - High school graduate, diploma, or equivalent (e.g., GED)
  - Some college, no degree
  - Trade/technical/vocational training
  - Associate degree
  - Bachelor’s degree
  - Master’s degree
  - Professional degree
  - Doctorate degree
  - Other: _______________
- What subject would you say was your strongest in school?
  - (Open response)
- If you attended a college or university, what was your major?

  - (Open response)
  - Did not attend
- What is your first language?
  - (Open response)
- Do you speak a second language?
  - No
  - Yes: _____________
- What do you consider yourself to be (choose all that apply)?
  - American Indian / First Nations
  - Asian
  - Native Hawaiian or other Pacific Islander
  - South American
  - Black / African American
  - Caucasian / European Descent
  - Other: ___________

Appendix C

Bem's Sex Role Inventory (BSRI) Used in Experiments 1-2

Please indicate on a 7-point scale how well each of the following personality characteristics describes yourself, ranging from 1 ("Never or almost never true") to 7 ("Always or almost always true"):

<u>Masculine items:</u>	<u>Feminine items:</u>	<u>Neutral items:</u>
Acts as a leader	Affectionate	Adaptable
Aggressive	Cheerful	Conceited
Ambitious	Childlike	Conscientious
Analytical	Compassionate	Conventional
Assertive	Does not use harsh language	Friendly
Athletic	Eager to soothe hurt feelings	Happy
Competitive	Feminine	Helpful
Defends own beliefs	Flatterable	Inefficient
Dominant	Gentle	Jealous
Forceful	Gullible	Likable
Has leadership abilities	Loves children	Moody
Independent	Loyal	Reliable
Individualistic	Sensitive to the needs of others	Secretive
Makes decisions easily	Shy	Sincere
Masculine	Soft spoken	Solemn
Self-reliant	Sympathetic	Tactful
Self-sufficient	Tender	Theatrical
Strong personality	Understanding	Truthful
Willing to take a stand	Warm	Unpredictable
Willing to take risks	Yielding	Unsystematic

Appendix D

Distractor Task for Experiments 1-3

Please solve as many of the following math problems as you can during the next two minutes. Your answers will not be evaluated

1. 7 x 3 =
2. 24 – 9 =
3. 18 + 23 =
4. 120 ÷ 6 =
5. 59 – 23 =
6. 11 x 4 =
7. 27 + 39 =
8. 12 – 6 =
9. 13 x 3 =
10. 64 ÷ 4 =
11. 21 + 6 =
12. 97 – 13 =
13. 6 x 8 =
14. 42 + 19 =
15. 81 ÷ 9 =
16. 27 – 13 =
17. 11 x 9 =
18. 19 + 21 =
19. 7 – 4 =
20. 10 x 8 =
21. 20 + 28 =
22. 25 – 20 =
23. 10 x 2 =
24. 24 ÷ 3 =
25. 13 + 16 =
26. 35 – 19 =
27. 14 x 8 =
28. 12 ÷ 6 =
29. 2 x 17 =
30. 37 – 14 =
31. 35 + 19 =
32. 42 ÷ 6 =
33. 24 – 3 =
34. 13 x 3 =
35. 16 ÷ 2 =
36. 22 + 10 =
37. 39 – 37 =
38. 13 x 2 =
39. 31 + 7 =
40. 48 ÷ 3 =
41. 50 – 32 =
42. 15 x 8 =
43. 6 + 4 =
44. 32 – 4 =
45. 24 ÷ 2 =
46. 20 x 13 =
47. 5 + 4 =
48. 15 x 5 =
49. 44 ÷ 11 =
50. 24 + 14 =
51. 32 – 17 =
52. 19 + 36 =
53. 39 – 1 =
54. 36 ÷ 9 =
55. 15 + 3 =
56. 16 – 3 =
57. 10 x 6 =
58. 18 ÷ 2 =
59. 9 – 6 =
60. 50 + 19 =
61. 36 ÷ 9 =
62. 5 x 5 =
63. 28 – 23 =
64. 45 ÷ 3 =
65. 10 + 3 =
66. 9 x 2 =
67. 18 – 4 =
68. 18 ÷ 3 =
69. 10 x 5 =
70. 37 + 12 =
71. 33 – 10 =
72. 42 ÷ 3 =
73. 25 + 5 =
74. 9 x 7 =
75. 26 – 24 =
76. 50 + 32 =
77. 49 ÷ 7 =
78. 7 x 8 =
79. 39 – 32 =
80. 20 + 7 =
81. 8 x 3 =
82. 37 – 12 =
83. 35 + 13
84. 42 ÷ 21 =
85. 17 x 10 =
86. 19 + 3 =
87. 6 ÷ 3 =
88. 4 x 3 =
89. 21 – 16 =
90. 25 + 2 =
91. 30 ÷ 5 =
92. 11 x 4 =
93. 13 + 4 =
94. 29 – 19 =
95. 46 + 6 =
96. 6 x 6 =
97. 8 x 7 =
98. 34 + 34 =
99. 10 x 5 =

Appendix E

Role Images used in Experiments 2 and 3

Architect

Boxer

Bricklayer

Carpenter

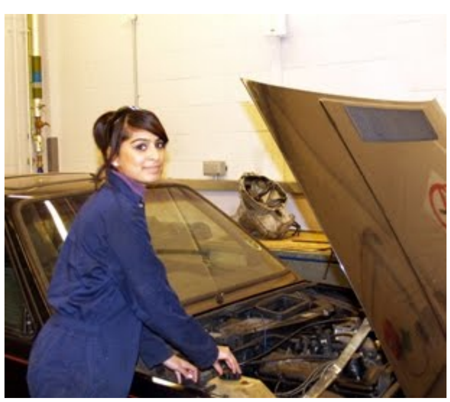
Electrician

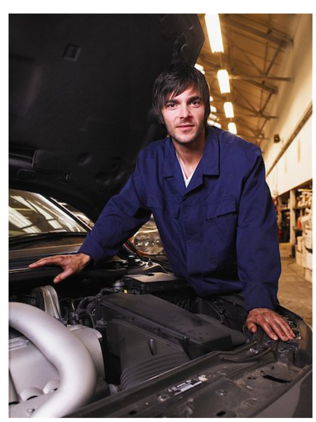
Farmer

Golfer

Judge

Mechanic





Soldier

Surgeon

Truck Driver

Au Pair


Ballet Dancer

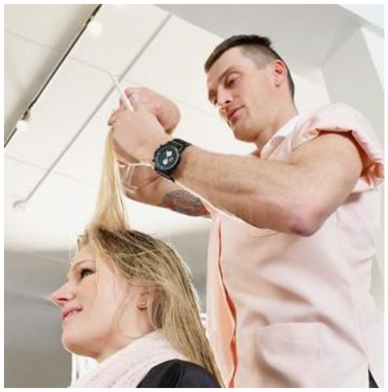

Cleaner


Flight Attendant

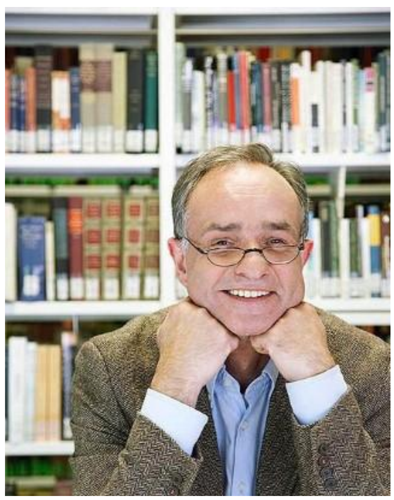
Florist

Fortune Teller

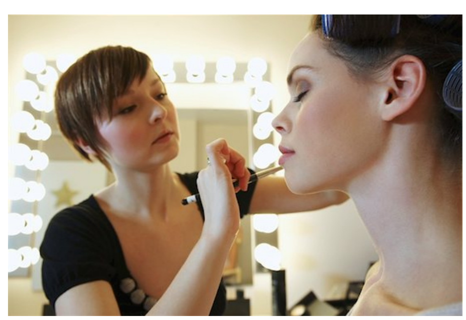
Hairdresser



Librarian

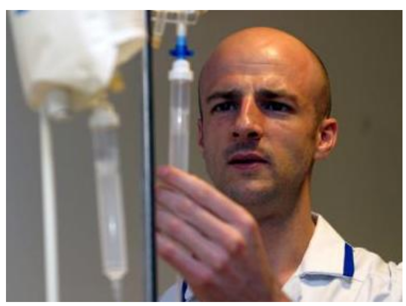


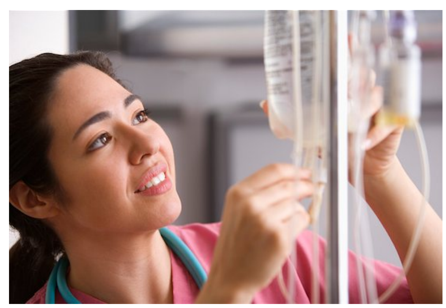
Make-Up Artist



Model

Nurse





Primary School Teacher

Appendix F

Mathematical Test (Practice ACT Exam Questions) Used in Experiment 3

---

1. $4x^3 \times 3xy^2 \times 2xy^2$ is equivalent to:
    a. $9x^3y^4$
    b. $9x^5y^4$
    c. $24x^3y^4$
    d. $24x^5y^4$
    e. $24x^5y^6$

2. Mr. Wilk is a high school math teacher whose salary is $33,660 for this school year, which as 180 days. In Mr. Wilk's school district, substitute teachers are paid $85 per day. If Mr. Wilk takes a day off without pay and a substitute teacher is paid to teach his classes, how much less does the school district pay in salary by paying a substitute teacher instead of Mr. Wilk for that day?
    a. $57
    b. $85
    c. $102
    d. $114
    e. $187

3. A student has earned the following scores on four 100-point tests this marking period: 63, 72, 88, and 91. What score must the student earn on the fifth and final 100-point test of the marking period to earn an average test grade of 80 for the five tests?
    a. 79
    b. 86
    c. 89
    d. 94
    e. The student cannot earn an average of 80.

4. The oxygen saturation of a lake is found by dividing the amount of dissolved oxygen the lake water currently has per liter by the dissolved oxygen capacity per liter of the water, and then converting that number into a percent. If the lake currently has 6.4 milligrams of dissolved oxygen per liter of water and the dissolved oxygen capacity is 9.5 milligrams per liter, what is the oxygen saturation level of the lake, to the nearest percent?
    a. 64%
    b. 67%
    c. 70%
    d. 89%
    e. 95%

5. A rectangular lot that measures 125 feet by 185 feet is completely fenced. What is the length, in feet, of the fence?
    a. 310
    b. 435
    c. 620
    d. 740
    e. 1,240

6. The expression $a[(b-c)+d]$ is equivalent to:
    a. $ab+ac+ad$
    b. $ab-ac+d$
    c. $ab-ac+ad$
    d. $ab-c+d$
    e. $a-c+d$

7. If $6x-3=-5x+7$, then x =?
    a. $\frac{4}{11}$
    b. $\frac{10}{11}$
    c. $\frac{11}{10}$
    d. $\frac{1}{2}$
    e. 10

8. What two numbers should be placed in the blanks below so that the difference between the consecutive numbers is the same?

    13, ___, ___, 34

    a. 19, 28
    b. 20, 27
    c. 21, 26
    d. 23, 24
    e. 24, 29

9. If *x* is a real number such that $x^3=729$, then $x^2+\sqrt{x}$ =?
    a. 9
    b. 27
    c. 30
    d. 84
    e. 90

10. If a gumball is randomly chosen from a bag that contains exactly 6 yellow gumballs, 5 green gumballs, and 4 red gumballs, what is the probability that the gumball chosen is NOT green?
    a. $\frac{2}{3}$
    b. $\frac{1}{3}$
    c. $\frac{2}{5}$
    d. $\frac{3}{5}$
    e. $\frac{4}{15}$

11. The number of students participating in fall sports at a certain high school can be shown with the following matrix:

Tennis	Soccer	Cross-Country	Football
25	30	50	80

The athletic director estimates the ratio of the number of sports awards that will be earned to the number of students participating with the following matrix:

Tennis	0.2
Soccer	0.5
Cross-Country	0.3
Football	0.4

Given these matrices, what is the athletic director's estimate for the number of sports awards that will be earned for these fall sports?
    a. 55
    b. 60
    c. 65
    d. 67
    e. 74

12. What expression must the center cell of the table below contain so that the sums of each row and each column are equivalent?

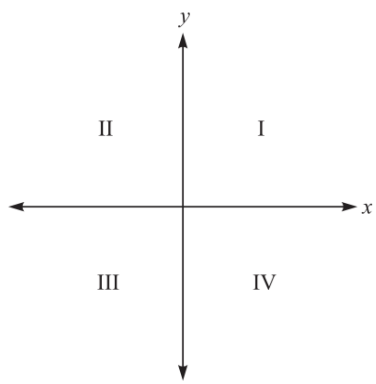
$-4x$	$9x$	$2x$
$7x$		$-3x$
$4x$	$-5x$	$8x$

   a. $5x$
   b. $3x$
   c. 0
   d. $-x$
   e. $-4x$

13. Reggie knows how to make 5 different entrees, 4 different side dishes, and 6 different desserts. How many distinct complete meals, each consisting of an entrée, a side dish, and a dessert, can Reggie make?
   a. 16
   b. 26
   c. 72
   d. 120
   e. 144

14. At a bottling plant, 10,000 liters of carbonated water are needed to produce 3,000 bottles of soda. How many liters of carbonated water are needed to produce 750 bottles of soda?
   a. 225
   b. 1,500
   c. 2,500
   d. 4,000
   e. 5,000

15. Point A is to be graphed in a quadrant, not on an axis, of the standard (x, y) coordinate plane below. If the x-coordinate and the y-coordinate of point A are to have the same signs, then point A *must* be located in:

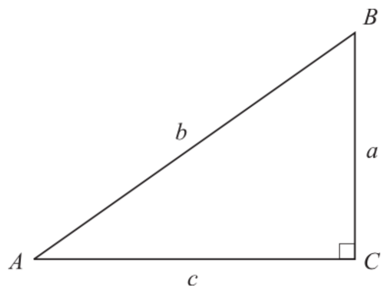


a. Quadrant I only
b. Quadrant II only
c. Quadrant III only
d. Quadrant I or II only
e. Quadrant I or III only

16. What is the slope-intercept form of $6x - 2y - 4 = 0$?
a. $y = 6x - 2$
b. $y = 3x + 2$
c. $y = 3x - 2$
d. $y = -3x + 2$
e. $y = -6x - 4$

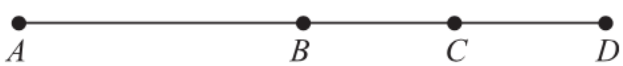
17. Which of the following is a solution to the equation $x^2 + 25x = 0$?
a. 50
b. 25
c. 5
d. -5
e. -25

18. For the right triangle ΔABC shown below, what is tan B?



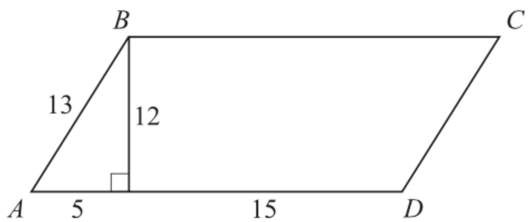
a. $\frac{a}{b}$
b. $\frac{a}{c}$
c. $\frac{b}{a}$
d. $\frac{c}{a}$
e. $\frac{c}{b}$

19. Points B and C lie on segment AD as shown below. The length of segment AD is 25 units; the segment AC is 19 units long; and the segment BD is 14 units long. How many units long, if it can be determined, is the segment BC?



a. 5
b. 6
c. 8
d. 11
e. Cannot be determined from the given information.

20. Parallelogram *ABCD*, with dimensions in inches, is shown in the diagram below. What is the area of the parallelogram, in square inches?



a. 60
b. 72
c. 180
d. 240
e. 260

Appendix G

Math Anxiety Survey Used in Experiment 3

For each statement circle a number 1-5 which indicates whether you strongly agree (5), agree (4), no opinion (3), disagree (2) or strongly disagree (1).

1. I usually have been at ease in math classes.	1 2 3 4 5
2. I see math as a subject I will rarely use.	1 2 3 4 5
3. I'm no good at math.	1 2 3 4 5
4. Generally, I have felt secure about attempting math.	1 2 3 4 5
5. I'll need mathematics for my future work.	1 2 3 4 5
6. I'd be happy to get good grades in mathematics.	1 2 3 4 5
7. I don't think that I could do advanced math.	1 2 3 4 5
8. It wouldn't bother me at all to take more math courses.	1 2 3 4 5
9. For some reason, even though I study, math seems unusually hard for me.	1 2 3 4 5
10. My mind goes blank and I am unable to think clearly when working in mathematics.	1 2 3 4 5
11. Knowing mathematics will help me earn a living.	1 2 3 4 5
12. Math has been my worst subject.	1 2 3 4 5
13. I think I could handle more difficult mathematics.	1 2 3 4 5
14. I'm not the type to do well in mathematics.	1 2 3 4 5
15. Math doesn't scare me at all.	1 2 3 4 5